Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 1515–1518, February 1994
Biochemistry

Design, creation, and characterization of a stable, monomeric triosephosphate isomerase

(loop deletion/monomerization/protein engineering)

TORBEN V. BORCHERT*, RUBEN ABAGYAN*, RAINER JAENICKE†, AND RIK K. WIERENGA*‡

*European Molecular Biology Laboratory, Postfach 102209, D-69012 Heidelberg, Germany; and †Institut für Biophysik und Physikalische Biochemie, Universität Regensburg, D-8400 Regensburg, Germany

Communicated by Hans Neurath, October 5, 1993

ABSTRACT **Protein engineering on trypanosomal triosephosphate isomerase (TIM) converted this oligomeric enzyme into a stable, monomeric protein that is enzymatically active. Wild-type TIM consists of two identical subunits that form a very tight dimer involving interactions of 32 residues of each subunit. By replacing 15 residues of the major interface loop by another 8-residue fragment, a variant was constructed that is a stable and monomeric protein with TIM activity. The length, sequence, and conformation of the designed fragment were suggested by extensive modeling.**

The $(\beta/\alpha)_8$ protein fold, commonly known as the triosephosphate isomerase (TIM) fold, seems to be nature's favorite choice of a scaffold for enzymatic activities as ≈20 different enzymes have, so far, been shown to have this fold (1, 2). The β-strands form an internal β-barrel consisting of 8 parallel β-strands (strands 1–8), covered on the outside by 8 α-helices (helices 1–8). The connections between β-strands and the subsequent α-helices are referred to as loops 1–8. The side chains of some of these loop residues form the active site of the TIM barrel enzymes and novel active sites can therefore theoretically be constructed at the C terminus of the β-strands without altering the overall fold. Our goal is to design activities on this $(\beta/\alpha)_8$ framework. Interestingly, most of the known $(\beta/\alpha)_8$ proteins are part of homo- or heterooligomers. For our studies, TIM from *Trypanosoma brucei* was chosen as the model protein because the crystal structure is known at 1.83 Å resolution (3). TIM forms very stable dimers with high catalytic activity. These properties have hindered the accurate measurements of the activity of individual monomers; the monomers have been postulated to be inactive in both renaturation (4, 5) and mutagenesis (6) studies, whereas matrix-bound monomers have been reported to retain activity (7). No allosteric control or cooperativity has been found between the two subunits (8, 9). The K_d of the TIM dimer is not known, but based on the low protein concentration in the activity assay, and assuming that the monomer is much less active than the dimer, then the K_d value can be estimated to be <10 pM. The first step in our protein engineering scheme has been to convert the dimeric TIM to a monomeric protein to (*i*) facilitate the subsequent design of active sites, (*ii*) examine the functional and structural importance of dimer formation, and (*iii*) have the principal possibility to obtain an active monomer. Only few such mutagenesis experiments of other proteins have been described. One example is the protein hormone insulin, in which it has been shown that single and double amino acid alterations of interface residues led to considerable variations in the association of the insulin monomers (10). The dimeric λCro DNA binding protein has been converted to a stable monomer that binds to DNA (11). Both cases involve small nonenzyme proteins on the order of 60 residues and insulin is, in fact, functional as a monomer. Single amino acid alterations of interface residues of the dimeric tyrosyl-tRNA synthetase resulted in inactive monomers (12).

The TIM dimer interface is very extensive, consisting mostly of residues from loops 1–4 (3). Each subunit buries ≈1600 Å^2 (3) surface area by dimerization. This buried area is expected to contribute significantly to protein stability (13). The active site is located close to the dimer interface, but the residues of one subunit do not interact directly with substrate analogues bound in the active site of the other subunit. In the crystal structure of the TIM–glycerol 3-phosphate complex (3) the closest distance is 5.4 Å between OG1 of Thr-75 (subunit 1) and the terminal carbon atom, C3, of glycerol 3-phosphate bound to subunit 2. Thr-75 is located at the tip of loop 3, which is a long loop (residues 65–79) protruding 13 Å out of the bulk of its subunit. This loop fits into a deep crevice near the active site of the other subunit (Fig. 1*a*) such that it stabilizes the active site architecture. Loop 3 atoms provide most of the stabilizing interactions at the dimer interface and many of these atoms are completely buried due to dimerization. In trypanosomal TIM, there are 256 intersubunit atom–atom pairs within a 4-Å cutoff (disregarding hydrogens); 203 of these involve loop 3 residues. In a similar manner, 17 of 20 interface hydrogen bonds involve loop 3 atoms (15). Point mutations preventing the association of individual subunits will result in TIM monomers with a deep pocket, a large protruding loop, and an extensive, solvent exposed, hydrophobic patch (Fig. 1) if structural rearrangements do not occur. Therefore, to obtain a monomeric and stable TIM variant, the deletion of loop 3, together with a reduction of the size of the hydrophobic patch, was investigated by careful modeling studies. The suggested residue changes were made and here we report on the physical, chemical, and catalytic properties of this loopless variant of trypanosomal TIM, which will be referred to as monotim.

MATERIALS AND METHODS

Modeling. All modeling operations (16, 17) and energy calculations were performed by the ICM program developed for efficient global optimization of varied energy and penalty functions in an arbitrarily fixed multimolecular system using the internal coordinates as variables. To design the new loop we created a smaller system composed of the loop 3 residues Gln-65–Asp-85 and the surrounding residues 45, 91–93, 98, and 112–121. The system was regularized and energy minimized with all hydrogens added. In the final modeling calculations, residue 65 and the surrounding fragments were kept fixed and residues 66 and 81–85 were tethered to their crystallographic positions; residues 67–80 were unrestrained.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TGGE, temperature gradient gel electrophoresis; TIM, triosephosphate isomerase.

‡To whom reprint requests should be addressed.

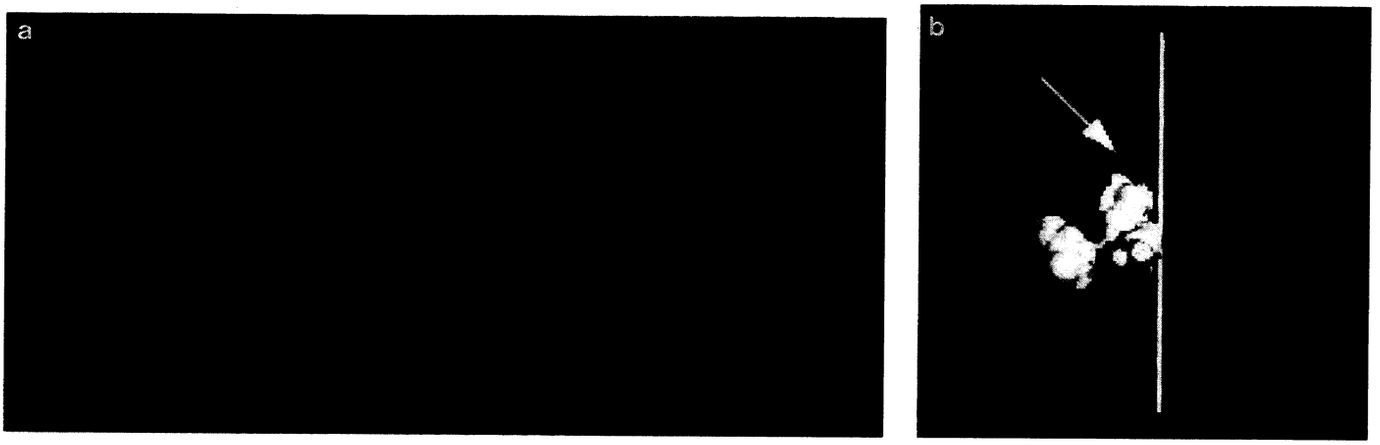

FIG. 1. Molecular surface (14) of the interface region of the trypanosomal TIM monomer. On the stereo pair (*a*), the hydrophobic atoms (C, S) are red, and the polar atoms (N, O) are purple. The monopicture (*b*) has the same view and color code; however, the loop 3 residues (residues 72–78) are white (Thr-75, located at the tip of loop 3 is yellow), Val-46 is blue, Cys-14 is green, and the atoms of Asn-11 (at the bottom of the active site pocket) are yellow. White line indicates position of the local twofold axis. Note the hydrophobic patch indicated by the white arrow. In the stereo pair the exposed Cys-14, Val-46, and Thr-75 clearly protrude out of the monomer. After applying the transformation of the local twofold axis, Cys-14 fits in a pocket constituted of loop 3 atoms, Val-46 contacts the hydrophobic patch, and the protruding Thr-75 moves into a deep pocket, which is connected with the active site pocket near Asn-11.

The ECEPP (Empirical Conformational Energy Program for Peptides) energy function (18) was combined with the solvation energy term (19). The global optimization was performed by the biased probability Monte Carlo procedure (17) applied to the torsion angles of the trial loop combined with minimization to achieve loop closure and relax the clashes. Side chain optimization was done by the same procedure applied only to the side chain torsion angles.

The closest C^{α} atoms in the loop (4.7 Å) were those of residues 68 and 79. Therefore, we started with glycine insertion variants between Ile-68 and Ser-79. Attempts to model a connection with one or two linking glycines between residues Ile-68 and Ser-79 resulted in a model with steric strain. For example, some of the $\phi\psi$ values of the models with the lowest energy were outside the acceptable regions of the Ramachandran plot. However, the substitution of Ala for Pro-81, and having three linking residues, allowed for an extra helical turn at the N terminus of helix 3, which could then smoothly be connected to residue 68. To obtain sufficient conformational flexibility, Ile-68 was changed into a glycine.

The overall hydrophobicity of the dimer interface is only slightly higher than the dimer outer surface (15), but side chains of Phe-45, Leu-48, Val-78, Ile-82, and Phe-86 form the previously mentioned hydrophobic patch at the interface of trypanosomal TIM. In the designed monotim, the overall hydrophobicity in this region is reduced because Val-78 is deleted; in addition, Ile-82 has been changed into Ser (Fig. 2*b*). The sequence and modeled structure of monotim are compared to wild-type TIM in Fig. 2.

Mutagenesis and Purification. Mutagenesis was carried out on plasmid pTIM (20) by the overlap extension method using the polymerase chain reaction (21). DNA primers were designed in such a way that mainly preferred *Escherichia coli* codons for highly expressed proteins (22) were used for inserting amino acids. The amplified DNA fragment, containing the mutations, was subcloned into pTIM and the sequence of the complete fragment was verified.

The soluble monotim, expressed in *E. coli* strain BL21(DE3) (23) grown in M9 medium (24) at 25°C and induced with 400 μM isopropyl β-D-thiogalactopyranoside was purified by the following procedure. The cells were lysed by sonication in 100 mM triethanolamine (TEA) HCl buffer (pH 8.0) containing 50 mM NaCl, 1 mM dithiothreitol (DTT), 1 mM NaN_3, and 0.5 mM phenylmethylsulfonyl fluoride. The supernatant from a precipitation step with 0.5% polyethylenimine was ammonium sulfate precipitated. The 50–70% saturation fraction was dialyzed against a 20 mM TEA HCl buffer (pH 8.0) supplemented with 20 mM NaCl and 1 mM each DTT, EDTA, and NaN_3. The soluble proteins were loaded on an S-Sepharose ion-exchange column (Pharmacia) and eluted with a 20–120 mM NaCl gradient in the same buffer as described above. This ion-exchange step makes it very unlikely that the purified monotim is contaminated with *E. coli* TIM, because the pI of trypanosomal TIM is 9.8 (25) and the pI of *E. coli* TIM is 5.3 (26). The yield was 6 mg of pure protein from 1 liter of culture.

Characterization. The stability of monotim was determined by temperature gradient gel electrophoresis (TGGE). This

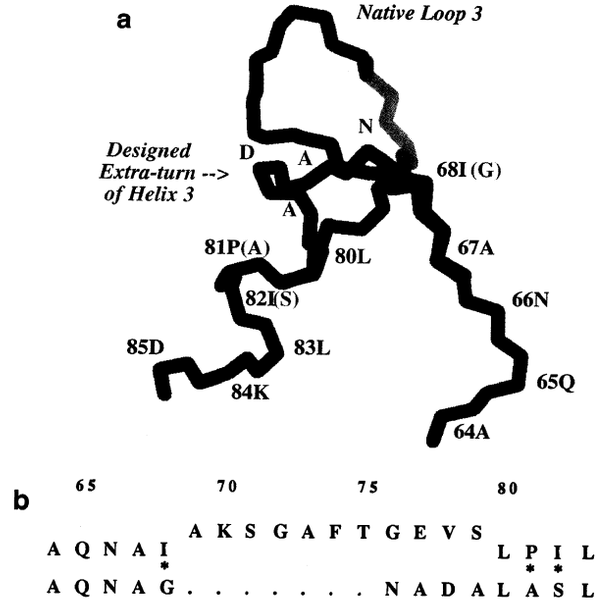

FIG. 2. (*a*) The N, C^{α}, C trace of the residues near loop 3 of wild-type TIM (purple) and the modeled monotim (yellow). (*b*) Alignment of amino acid (one-letter code) sequences of wild-type *T. brucei* TIM (top line) and the designed monotim (bottom line). In wild-type TIM Ala-64 is the last residue of β-strand 3 and Leu-80 is the first residue of helix 3. Residues 69–79 are exchanged with a 4-amino acid additional turn of helix 3. Additional amino acid alterations of flanking residues are indicated by asterisks.

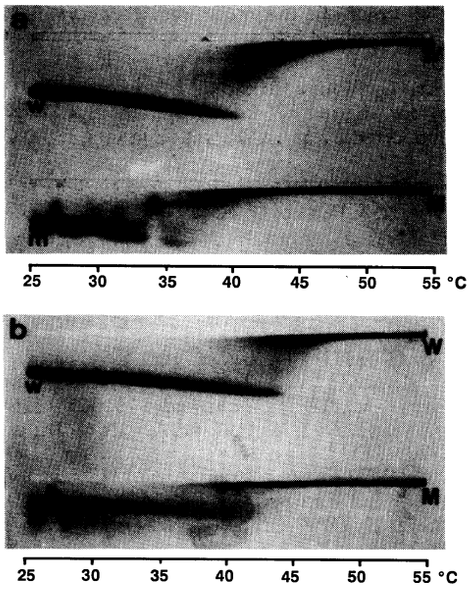

FIG. 3. Thermal stability of wild-type TIM and monotim as measured by TGGE. TGGE experiments with monotim and wild-type TIM were carried out in the absence (*a*) or presence (*b*) of 0.5 mM 2-phosphoglycollate. The t_m of wild-type TIM is 41°C in the absence of 2-phosphoglycollate and ≈44°C in the presence of the inhibitor. Monotim has a t_m of 38°C without 2-phosphoglycollate; the transition is somewhat broader in the presence of 2-phosphoglycollate—namely, between 38°C and 43°C. The temperature gradient is shown below the protein bands. W, aggregated wild-type TIM; M, aggregated monotim. The band of migrating, native protein is indicated as follows: w, wild-type TIM; m, monotim.

TGGE system (Diagen GmbH) is a native gel with a temperature gradient perpendicular to the electric field (27). The gels contain 100 mM Bistris acetate (pH 7.0) and 8% acrylamide/bisacrylamide (29:1). A 100-μg sample of the native protein, loaded horizontally across the gel, migrates into the gel at constant temperature (20°C). A linear temperature gradient from 25°C to 55°C, as indicated, is formed and the gel run is continued. At temperatures above melting temperature (t_m), the proteins denature and aggregate. The aggregated proteins do not migrate in the gel. The proteins are stained with Coomassie blue.

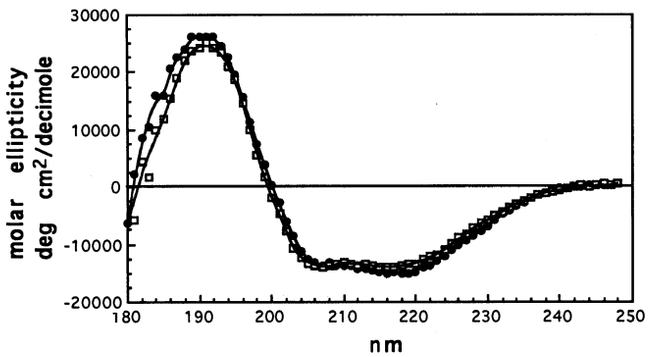

FIG. 4. Molar ellipticity was measured for wild-type TIM (●) and monotim (□) between 180 and 250 nm. Two conclusions can be drawn from these data. First, the monotim spectrum shows a high percentage of secondary structure. Second, the similarity to the wild-type TIM spectrum indicates that the two proteins have approximately the same kind of secondary structure. The subtle differences between the spectra could be due to small changes in the secondary structure or to the different environments of some aromatic side chains, especially certain phenylalanines. Loop 3 residue Phe-74 is not present in monotim and the former interface residues Phe-45, Phe-86, Tyr-101, and Tyr-102 will be more exposed to solvent in monotim.

The CD spectra were recorded on a Jobin Yvon circular dichrometer VI with a 0.2-mm pathlength at 20°C. For these experiments, solutions of wild-type TIM and monotim were equilibrated against a 50 mM sodium phosphate buffer (pH 7.0) supplemented with 20 mM NaCl. Samples were diluted with the same buffer to a concentration of 4.5 μM. Concentrations were determined by measuring absorbance at 280 nm.

Gel-filtration studies were carried out with a column (45 × 0.79 cm^2) of Pharmacia Sephacryl S200 matrix. The column was equilibrated and eluted with 100 mM TEA HCl buffer (pH 8.0) including 500 mM NaCl, 1 mM DTT, 1 mM NaN_3, and 1 mM EDTA. Monotim (300 μg) was loaded on the column. The concentration of protein in the fractions was determined by absorbance at 280 nm with a 1-cm pathlength. A standard curve of partition coefficient, $K_{av} = (v_e - v_o)/(v_t - v_o)$, versus molecular mass was established with blue dextran 2000 and proteins of 14, 25, 43, and 67 kDa. TIM activity was measured in the peak fractions as described (28).

The ultracentrifugation experiments (29) were performed in a Beckman Spinco analytical ultracentrifuge (model E) equipped with a UV-scanning system. Double sector cells with sapphire windows were used in AuF-Ti and AuG rotors. The temperature was 20°C ± 2°C. *s* values were determined at 50,000 rpm, plotting ln*r* vs. time and correcting to 20°C and water viscosity. Sedimentation equilibria were evaluated from ln*c* vs. r^2 plots. The partial specific volume was calculated from the amino acid composition.

RESULTS AND DISCUSSION

Several physical, chemical, and kinetic properties of purified monotim have been determined. The stability of monotim was assayed by TGGE. This experiment (Fig. 3*a*) shows that the t_ms for wild-type TIM and monotim are 41°C and 38°C, respectively, which proves that monotim is approximately as stable as the wild-type protein. A TGGE experiment was also carried out in the presence of 0.5 mM 2-phosphoglycollate (Fig. 3*b*), which is a competitive inhibitor of TIM (K_i = 27 μM) (24). The stabilization of monotim as well as wild-type TIM by 2-phosphoglycollate indicates that monotim binds the substrate analogue, suggesting that the TIM active site topography is conserved in monotim. The similarity of the far UV CD spectra of monotim and wild-type TIM (Fig. 4) further indicates that monotim also has the $(\beta/\alpha)_8$ fold.

The molecular mass of monotim was determined by gel filtration to be 22 kDa (Fig. 5). Sedimentation analysis was applied to more precisely determine the size of monotim at a protein concentration ranging from 0.02 to 2 mg/ml. As taken from sedimentation velocity runs, the enzyme migrates as one symmetrical band with a sedimentation coefficient $s_{20,w}$ = 2.48 ± 0.05 S, practically independent of the concentration (0.2–2.0 mg/ml). High-speed sedimentation equilibrium experiments at various concentrations show a homogeneous monomer of 25,360 ± 980 Da, in agreement with the calculated molecular mass of 26,063 Da. Even in the bottom region of the cell, no tendency to form dimers or higher aggregates is detectable. Monotim is, indeed, a monomeric protein.

The observed catalytic activity of monotim, as shown in Fig. 5, is considerable, which is very surprising, but it is significantly reduced compared to the wild-type enzyme. It is very unlikely that it is due to a contamination with *E. coli* TIM, because in the gel-filtration peak the specific activity is constant (Fig. 5). The observed enzymatic activity is probably the best proof that the overall structure of monotim is very similar to wild-type TIM. The kinetic parameters were determined by varying the concentration of glyceraldehyde 3-phosphate in the enzyme assay (28) between 0.5 and 10.2 mM. K_m and k_{cat} values were determined by an Eadie–Hofstee plot to be 5.3 ± 0.9 mM and 312 min^{-1}, respectively.

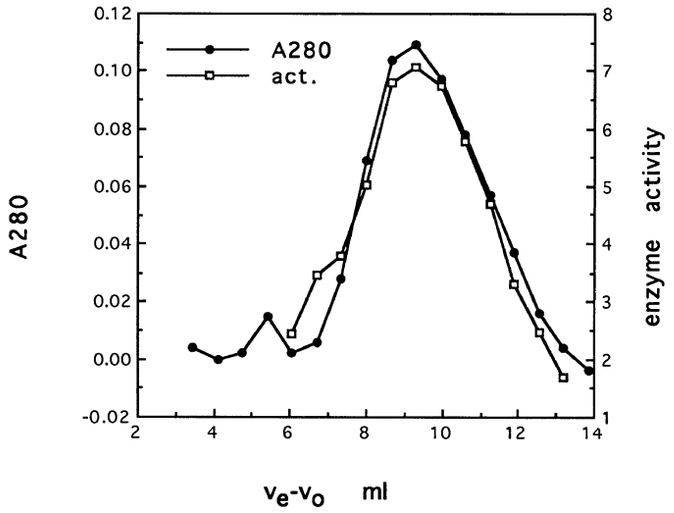

FIG. 5. Gel filtration of monotim. Elution profile (●) of monotim in a gel-filtration experiment is shown as well as the measured enzyme activity in the eluted fractions (□). At the *y* axis is plotted absorbance at 280 nm and enzymatic activity. Enzyme activity is given as nmol of converted D-glyceraldehyde 3-phosphate per min in a 100-μl sample. At the *x* axis is plotted $v_e - v_o$, representing elution volume minus void volume in ml. The partition coefficient, K_{av}, for monotim was 0.42, which corresponds to a molecular mass of 22 kDa. The identical experiment with wild-type *T. brucei* TIM resulted in a K_{av} of 0.27 (based on a $v_e - v_o$ value of 6 ml) and a deduced molecular mass of 50 kDa (data not shown). The calculated molecular masses for dimeric wild-type TIM and monotim are 54 and 26 kDa, respectively. The values obtained by this experiment are in excellent agreement with the monomeric appearance of monotim. The correlation of enzyme activity with the protein concentration in the eluted fractions showed that the TIM activity was caused by monotim and rules out contamination by *E. coli* TIM.

Compared to wild-type TIM, monotim has ≈20-fold lower affinity and 1000-fold lower k_{cat} for this substrate. The activity is not due to low-affinity dimerization as the specific activity is independent of protein concentration (Fig. 5).

This study has shown the successful design of major alterations of TIM leading to a monomeric TIM with considerable activity and provides an example of the conversion of an oligomeric enzyme into a stable monomeric protein that is still enzymatically active. Previously we have tried to reduce the stability of the TIM dimer by single amino acid substitutions of interface residues—e.g., the replacement of His-47 by an Asn (20). For this variant of TIM (H47N TIM) the dimer stability is reduced significantly compared to the wild type but dimers do form at high protein concentrations; the thermostability (t_m = 33°C by TGGE) of H47N TIM is also significantly lower (ref. 20; T.V.B., R.J., and R.K.W., unpublished observations). However, the replacement of loop 3 by a strainless turn concomitantly with the decrease of the hydrophobic patch, as suggested by the modeling studies, creates a stable, monomeric protein. We hope to determine the crystal structure of monotim, after which the design work aiming toward specific activities on this monomeric $(\beta/\alpha)_8$ scaffold can continue.

Note Added in Proof. In the meantime, the crystal structure of monotim has been solved (30).

We thank Dr. G. Böhme for the ultracentrifugation software and Dr. J. Lakey for sharing his knowledge about CD spectra. This research is supported by an EC-BRIDGE grant (BIOT-CT90-0182).

1. Brändén, C.-I. (1991) *Curr. Opin. Struct. Biol.* **1,** 978–983.
2. Farber, G. K. & Petsko, G. A. (1990) *Trends Biochem. Sci.* **15,** 228–234.
3. Wierenga, R. K., Noble, M. E. M., Vriend, G., Nauche, S. & Hol, W. G. J. (1991) *J. Mol. Biol.* **220,** 995–1015.
4. Zabori, S., Rudolph, R. & Jaenicke, R. (1980) *Z. Naturforsch.* **35,** 999–1004.
5. Waley, S. G. (1973) *Biochem. J.* **135,** 165–172.
6. Casal, J. I., Ahern, T. J., Davenport, R. C., Petsko, G. A. & Klibanov, A. M. (1987) *Biochemistry* **26,** 1258–1264.
7. Branford-White, C. J. & Fell, D. A. (1976) *Trans. Biochem. Soc.* **4,** 620–622.
8. Fersht, A. (1985) *Enzyme Structure and Mechanism* (Freeman, New York), 2nd Ed., pp. 439–440.
9. Schnackerz, K. D. & Gracy, R. W. (1991) *Eur. J. Biochem.* **199,** 231–238.
10. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Snel, L., Sörensen, A. R. & Vogt, H. O. (1988) *Nature (London)* **333,** 679–682.
11. Mossing, M. C. & Sauer, R. T. (1990) *Science* **250,** 1712–1715.
12. Jones, D. H., McMillan, A. J. & Fersht, A. R. (1985) *Biochemistry* **24,** 5852–5857.
13. Janin, J., Miller, S. & Chothia, C. (1988) *J. Mol. Biol.* **204,** 155–164.
14. Connolly, M. L. (1985) *J. Am. Chem. Soc.* **107,** 1118–1124.
15. Wierenga, R. K., Noble, M. E. M. & Davenport, R. C. (1992) *J. Mol. Biol.* **224,** 1115–1126.
16. Abagyan, R. A. & Argos, P. (1992) *J. Mol. Biol.* **225,** 519–532.
17. Abagyan, R. A. & Totrov, M. M. (1994) *J. Mol. Biol.* **235,** in press.
18. Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. (1975) *J. Phys. Chem.* **79,** 2361–2381.
19. Wesson, L. & Eisenberg, D. (1992) *Protein Sci.* **1,** 227–235.
20. Borchert, T. V., Pratt, K., Zeelen, J. P., Callens, M., Noble, M. E. M., Opperdoes, F. R., Michels, P. A. M. & Wierenga, R. K. (1993) *Eur. J. Biochem.* **211,** 703–710.
21. Higuchi, R., Krummel, B. & Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367.
22. Gribskov, M., Devereux, J. & Burgess, R. R. (1984) *Nucleic Acids Res.* **12,** 539–549.
23. Studier, F. W. & Moffatt, B. A. (1986) *J. Mol. Biol.* **189,** 113–130.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
25. Misset, O., Bos, O. J. M. & Opperdoes, F. R. (1986) *Eur. J. Biochem.* **157,** 441–453.
26. Mainfroid, V., Goraj, K., Rentier-Delrue, F., Houbrechts, A., Loiseau, A., Gohimont, A.-C., Noble, M. E. M., Borchert, T. V., Wierenga, R. K. & Martial, J. A. (1993) *Protein Eng.* **6,** 893–900.
27. Rosenbaum, V. & Riesner, D. (1987) *Biophys. Chem.* **26,** 235–246.
28. Lambeir, A.-M., Opperdoes, F. R. & Wierenga, R. K. (1987) *Eur. J. Biochem.* **168,** 69–74.
29. Yphantis, D. A. (1964) *Biochemistry* **3,** 297–317.
30. Borchert, T. V., Abagyan, R., Radha Kishan, K. V., Zeelen, J. P. & Wierenga, R. K. (1993) *Structure* **1,** 205–213.